

## The role of temporal cortex in the control of attention

Hamidreza Ramezanzpour<sup>a,b,c,\*</sup>, Mazyar Fallah<sup>a,b,c,d,e,\*\*</sup>

<sup>a</sup> Centre for Vision Research, York University, Toronto, Ontario, Canada

<sup>b</sup> School of Kinesiology and Health Science, Faculty of Health, York University, Toronto, Ontario, Canada

<sup>c</sup> VISTA: Vision Science to Application, York University, Toronto, Ontario, Canada

<sup>d</sup> Department of Psychology, Faculty of Health, York University, Toronto, Ontario, Canada

<sup>e</sup> Department of Human Health and Nutritional Sciences, College of Biological Science, University of Guelph, Guelph, Ontario, Canada

### ARTICLE INFO

#### Keywords:

Attention  
Temporal cortex attention network  
Executive control  
Active vision  
Cognitive programs  
Social interactions

### ABSTRACT

Attention is an indispensable component of active vision. Contrary to the widely accepted notion that temporal cortex processing primarily focuses on passive object recognition, a series of very recent studies emphasize the role of temporal cortex structures, specifically the superior temporal sulcus (STS) and inferotemporal (IT) cortex, in guiding attention and implementing cognitive programs relevant for behavioral tasks. The goal of this theoretical paper is to advance the hypothesis that the temporal cortex attention network (TAN) entails necessary components to actively participate in attentional control in a flexible task-dependent manner. First, we will briefly discuss the general architecture of the temporal cortex with a focus on the STS and IT cortex of monkeys and their modulation with attention. Then we will review evidence from behavioral and neurophysiological studies that support their guidance of attention in the presence of cognitive control signals. Next, we propose a mechanistic framework for executive control of attention in the temporal cortex. Finally, we summarize the role of temporal cortex in implementing cognitive programs and discuss how they contribute to the dynamic nature of visual attention to ensure flexible behavior.

### 1. Visual attention

The brain can only process a limited amount of information received by our sensory system at a given point in time. Attention selects behaviorally relevant information to overcome this bottleneck. Numerous studies have investigated neural and behavioral effects of visual attention in humans and nonhuman primates (Bichot and Schall, 1999, 2002; Corbetta and Shulman, 2002; Desimone and Duncan, 1995; Fallah et al., 2007; Reynolds et al., 1999; Schall and Hanes, 1993; Wang et al., 2015a).

Attention can be allocated either toward a specific spatial location (spatial attention), toward non-spatial features (feature-based attention) such as motion direction, color, or towards an object as defined by a combination of features at a location in the visual scene (object-based attention). Each of these types of attention has been shown to influence encoding of task-relevant locations or features throughout the visual cortical hierarchy (Cohen and Maunsell, 2011; Connor et al., 1997; Ipata

et al., 2012; Martinez-Trujillo and Treue, 2004; Maunsell and Treue, 2006; Motter, 1993) as well as higher-order areas such as the lateral intraparietal (LIP) area (Bisley and Goldberg, 2010; Ibos and Freedman, 2014, 2016), frontal eye field (FEF) (Armstrong et al., 2009), and lateral prefrontal cortex (LPFC; see Fig. 1) (Gaillard et al., 2020; Lennert and Martinez-Trujillo, 2011).

The premotor theory of attention postulated that the neural networks involved in eye movement control (the oculomotor system) and attentional control do not differ and visual attention is a consequence of action planning (Rizzolatti et al., 1987). This theory has received support from electrophysiological studies on monkeys (Ekstrom et al., 2008; Ignashchenkova et al., 2004; Kustov and Robinson, 1996; Lowet et al., 2018; Moore et al., 2003; Moore and Armstrong, 2003; Moore and Fallah, 2001) as well as neuroimaging studies in humans (Corbetta et al., 1998; Nobre et al., 2000). The most substantial support comes from a series of work on the monkey FEF, a critical control area for the oculomotor system, which showed that subthreshold stimulation of a motor

*Abbreviations:* TAN, Temporal cortex attention network; VAN, Ventral attention network; DAN, Dorsal attention network; GFP, Gaze following patch; pITd, Dorsal part of the posterior inferotemporal cortex; mid-STs, Middle part of the superior temporal sulcus.

\* Corresponding author. Centre for Vision Research, York University, Toronto, Ontario, Canada.

\*\* Corresponding author. Department of Human Health and Nutritional Sciences, College of Biological Science, University of Guelph, Guelph, Ontario, Canada.

*E-mail addresses:* [hamidram@yorku.ca](mailto:hamidram@yorku.ca) (H. Ramezanzpour), [mfallah@uoguelph.ca](mailto:mfallah@uoguelph.ca) (M. Fallah).

<https://doi.org/10.1016/j.crneur.2022.100038>

Received 27 November 2021; Received in revised form 5 February 2022; Accepted 1 April 2022

Available online 12 April 2022

2665-945X/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

vector, while not evoking a saccade, deploys attention to the location in the visual field that represents the endpoint of the motor plan (Moore et al., 2003; Moore and Armstrong, 2003; Moore and Fallah, 2001). This was followed up by a similar study in the superior colliculus (SC), a subcortical oculomotor area (Ignashchenkova et al., 2004). Notwithstanding these pieces of evidence, some other studies have dissociated the coupling between endogenous attention and eye movements (Hanning et al., 2019; Smith and Schenk, 2012). The limitation of premotor theory is that it is action-based and primarily driven by external stimuli.

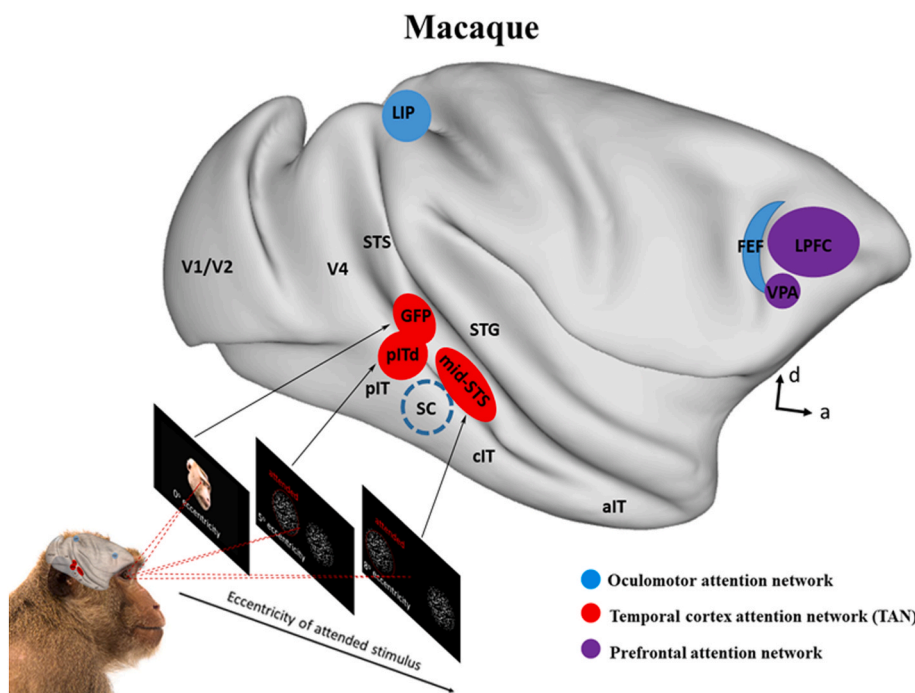
In addition to the premotor theory of attention, several other theories (Desimone and Duncan, 1995; Humphreys et al., 1998; Itti and Koch, 2001; Reynolds and Heeger, 2009; Fiebelkorn and Kastner, 2019; Bundesen et al., 2005; Olshausen et al., 1993; Tsotsos et al., 1995) have tried to encompass broader aspects of visual attention including both spatial and non-spatial (feature-based) attention. A detailed description of all of those models is beyond the scope of this review. Nevertheless, here we briefly review the concept of a few of these models which are relevant for temporal cortex participation in attentional control.

The selective tuning model is based on the brain’s hierarchical organization and assumes three stages: first, the stimuli enter the first layer and propagate to the upper layers via feed-forward connections in an inverted sub-pyramid manner. Second, a “winner take all” (Koch and Ullman, 1985) process is applied to each layer of the network, starting from the output layer and back-propagating towards the input layer. Consequently, at each layer, the connections that are not contributing to the winner are pruned away. The pruned connections form a suppressive annular region around the connections that remain active and form the attention zone that selects parts of the stimuli (Boehler et al., 2009; Yoo et al., 2018). Finally, in the third stage, the selected parts of the stimuli in the input layer propagate once again towards the output layer, but this time as if there are no distractors (Tsotsos et al., 1995). The selective tuning model is of interest here because it will enable us to explain the nature of temporal cortex participation in attentional control and how it might depend on the stimulus position in the visual field. The selective tuning model also fits with the findings of the recent studies on the temporal cortex (Bogadhi et al., 2018, 2019; Ramezanzpour and Thier, 2020; Stemmann and Freiwald, 2016, 2019) and supports the integration of cognitive programs which will be discussed in the next sections,

but that does not automatically discount other models of visual attention. Another computational model which is able to explain seemingly disparate experimental findings in the domain of visual attention is the normalization model of attention (Reynolds and Heeger, 2009). In this model, neural activity (designated as “stimulus drive”) is integrated with an external “attention field” and a “suppressive field”, that pools responses to non-preferred stimuli and unattended locations, which is used as in normalization. According to this model, attention modulates the strength of normalization which appears to operate at all stages of the visual system regardless of what biophysical mechanism it entails.

Many of these models of selective attention are dependent upon saliency and feature conspicuity maps. It has been hypothesized every visual scene can be segmented into separate feature conspicuity maps (Koch and Ullman, 1985; Treisman and Gelade, 1980) representing a single feature, such as orientation or color (Itti and Koch, 2001; Koch and Ullman, 1985; Fecteau and Munoz, 2006). These feature conspicuity maps are topographically organized and compete for selection. Topographic representation of the weighted sums of feature conspicuity map activations may generate a single master map (or “central representation” (Koch and Ullman, 1985)) representing saliency in a passive (bottom-up) manner (Koch and Ullman, 1985; Itti et al., 1998). The observer’s goal does not have any role in this type of processing. The saliency map is a theoretical framework, and its neural correlate (if it exists at all) remains an open question. Nevertheless, studies have tried to find the locus of the saliency map in the brain (Gottlieb et al., 1998; Thompson et al., 1997; Zenon et al., 2010; Li, 2002; Lee et al., 1999; Robinson and Petersen, 1992; Bisley and Goldberg, 2003; Goldberg et al., 2002; Gottlieb, 2002; McPeck and Keller, 2002; Noudoost et al., 2014; Thompson and Bichot, 2005). While the concept of saliency and conspicuity maps can help us understand many aspects of visual attention, behavior is best represented by a combination of the feature conspicuity maps with top-down relevance signals which reflect an active observer’s goal. Hitherto, studies have suggested that these integrations mostly happen in oculomotor structures, i.e., FEF, SC, LIP (see ref (Fecteau and Munoz, 2006) for review).

Relatedly, a theory of visual attention (Bundesen, 1990) proposes that each stimulus that reaches short-term memory earlier than others can be represented and recognized later. Each stimulus’ individual



**Fig. 1.** Oculomotor, prefrontal and temporal cortex areas involved in the control of attention. Blue: The oculomotor system priority maps (Fecteau and Munoz, 2006), red: The temporal cortex attention network (TAN) (Bogadhi et al., 2019; Ramezanzpour and Thier, 2020; Stemmann and Freiwald, 2019), and purple: the prefrontal control areas (Bichot et al., 2015; Buschman and Miller, 2007). The presentation of the visual stimuli that are required to be attended at various eccentric locations might have shifted the locus of attention control signals in the posterior-middle temporal cortex. Note that the locus of priority maps are approximate locations based on coordinates found in the corresponding original papers. To confirm the exact relationship between attentional foci in the temporal cortex and the type and eccentricity of stimuli, future studies should carry out mapping these areas in the same animals. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

processing speed is affected by its attentional weight which, in turn, are affected by the task relevance and saliency of the stimulus' features (Bundesen et al., 2005; Bundesen, 1990). Bundesen and colleagues proposed that dynamic remapping of receptive fields of cortical neurons is the basis for setting the weights which are used for reallocation of attention i.e. the more neurons allocated to an object the higher the attentional weight of the object (Bundesen et al., 2005). The neural implementation of this model requires two successively computed forms of object representation (Bundesen, 1990), an elementary visual feature representation which is not accessible for goal directed actions (also called proto-objects (Schneider, 2013)), and a visual working memory representation (also called visual tokens (Schneider, 2013)). The various models of visual attention just mentioned are generally mutually inconsistent even though each gives a supportable perspective on the overall attention problem. In this paper, we review recent findings which suggest that temporal cortex regions act on feature conspicuity representations, contribute to guiding attention, and implement cognitive control signals.

## 2. Visual processing in the STS and IT cortex

The temporal cortex can be coarsely divided into four sub-regions: medial temporal cortex (MTC), superior temporal gyrus (STG), superior temporal sulcus (STS) and inferotemporal (IT) cortex (see (Gross, 1994) for a review). In this review, we are focusing on the STS and IT, as the MTC and STG have yet to evidence support for attentional control. Classical studies have considered two separate but parallel streams for processing of visual information: the dorsal stream (where pathway) dealing with spatial aspects of stimuli, and the ventral stream (what pathway) implicated in recognition of object (Mishkin and Ungerleider, 1982; Mishkin et al., 1983). Later theories emphasized that the dorsal stream information is used to guide actions while ventral stream processing is necessary for perception (Goodale and Milner, 1992). For this review, we are focusing on the ventral stream which includes areas V1, V2, V4, and IT cortex (TEO and TE in nonhuman primate studies). Processing of the visual input becomes progressively more complex along the ventral stream, taking form and color information to produce object recognition (DiCarlo et al., 2012; Koida and Komatsu, 2007; Kravitz et al., 2013; Logothetis et al., 1995).

IT cortex is coarsely divided into three subregions: posterior IT (pIT), central IT (cIT) and anterior IT (aIT) (Conway, 2018; Tanaka, 1996). Based on cytoarchitectural divisions, pIT corresponds to TEO while cIT and aIT correspond to TE (Conway, 2018). Early neurophysiological studies investigated stimulus selectivity and receptive fields properties of IT cortex neurons (Desimone et al., 1984; Gross et al., 1969), determining that IT neurons have large bilateral receptive fields and complex object selectivity, including faces. While single-unit recordings showed that there is a columnar organization for shape processing (Fujita et al., 1992) in the IT cortex, later neurophysiological and neuroimaging studies led to the discovery of several patches within IT cortex selective for specific categories of objects such as faces, houses, or more complex stimuli including scenes (Downing et al., 2001; Freiwald and Tsao, 2010; Lafer-Sousa and Conway, 2013; Tsao et al., 2006, 2008). Neuroimaging studies in humans also revealed other potential principles of the IT cortex, namely retinotopic organization and real-world size representation of objects (Konkle and Oliva, 2012; Levy et al., 2001). However, until very recently, the general principle governing IT cortex organization was unclear. In a comprehensive study combining fMRI, electrophysiology, microstimulation, and deep networks, Bao and colleagues put forward a unified theory of IT cortex organization by showing that monkey IT cortex is topographically organized into a map of low dimensional object space that is repeated three times with increasing invariance at each stage (Bao et al., 2020).

In contrast, the STS, which sits between the ventral and dorsal streams, is involved in a wide variety of functions (Deen et al., 2015; Hein and Knight, 2008), such as motion processing (Pelphrey et al.,

2003; Saito et al., 1986; Thompson et al., 2005), speech processing (Binder et al., 2000), audiovisual integration (Beauchamp et al., 2004; van Atteveldt et al., 2004), multisensory perception (Dahl et al., 2009; Hikosaka et al., 1988), and social interaction processing (Isik et al., 2017; Ong et al., 2020; Saxe and Kanwisher, 2003). While the exact correspondence between various anatomical regions of the monkey and the human STS has not yet been established, most of the above-mentioned functions are shared between the two species. While it has been shown that complex information processing in the STS is partially handled by a number of specialized modules (Deen et al., 2015), this does not preclude the fact that the multifunctionality of the STS might be due to its coactivation with distinct neural networks implicated in distinct tasks (Hein and Knight, 2008). The latter notion gets further support from the massive bidirectional connections of the STS with a range of higher-order brain areas such as the ventral and medial frontal cortex, lateral prefrontal and premotor areas, the parietal cortex, and mesial temporal regions (Seltzer and Pandya, 1989, 1994), strategically placing it as a functional link between early visual areas and higher order areas. For example, integrating different modalities of information, such as vision and audition, is highly beneficial for disambiguating social decision processes (Kraemer et al., 2020). Another example is facial expression recognition which builds upon the ability to combine biological motion and facial information (Fisher and Freiwald, 2015). In summary, the STS shows a wide functionality around integrating and linking complex information from early sensory cortices to be later processed by higher order areas.

## 3. STS and IT cortex: recipients of visual attention

Although some studies have shown that the effects of attention are widespread and impact most of the visual areas starting from the lateral geniculate nucleus (LGN) to the higher order temporal cortex areas such as TEO (O'Connor et al., 2002), the influence of attention on the functioning of the temporal cortex has been less elucidated than for early visual areas such as V2 and V4. Nevertheless, attention has been shown to modulate shape processing responses in the human STS (Corbetta et al., 1990). For example, when dynamic aspects of human faces such as gaze or expressions are attended, neural responses of the pSTS face responsive areas are modulated (Dobs et al., 2018; Hoffman and Haxby, 2000; Narumoto et al., 2001). One of the seminal studies of how attention influences the responses of ventral stream neurons was carried out by Moran and Desimone (1985). They recorded the response of neurons in areas V4 and the IT cortex of monkeys while presenting two visual stimuli simultaneously inside the neuron's receptive field. They found that the response to the unattended stimulus was dramatically reduced. Similar results were later reported in areas such as V2, V4, MT, MST and IT (Reynolds et al., 1999; Kastner et al., 1998; Treue and Maunsell, 1996). Visual attention was shown to increase neuronal responses without changing selectivity, which is described as gain modulation (Motter, 1993). Around the same time, other studies showed that attention not only affects neural responses but also improves behavioral performance during visual discrimination or visual search tasks (Moran and Desimone, 1985; Chelazzi et al., 1993; Richmond and Sato, 1987; Spitzer et al., 1988).

As visual areas were shown to be affected by visuospatial attention, further studies looked to determine where the attentional signals originated from. Studies have suggested that the prefrontal cortex is the primary driver of the temporal cortex during object recognition (Fuster et al., 1985; Rees et al., 1997; Wilson et al., 1993). In the search for the source of attention control signals in the prefrontal cortex that drive object recognition specifically, two studies emphasized the role of inferior frontal junction (IFJ) in humans and its homolog in monkeys, ventral pre-arcuate gyrus (VPA) (see Fig. 1). Using fMRI simultaneously with magnetoencephalography (MEG) in humans, Baldauf and Desimone tested the hypothesis that attention to different object categories may synchronize areas representing those categories in the temporal

cortex and higher-order areas in the prefrontal cortex (Baldauf and Desimone, 2014). They used a sequence of stimuli consisting of two object categories superimposed (faces and houses), fading in and out of a phase-scrambled noise mask at different frequencies. Depending on which object was attended, an area close to Brodmann areas 45 and 46, the IFJ, was found to be synchronously activated with the fusiform face area (FFA) and parahippocampus place area (PPA) which represent faces and places respectively. The phase analysis showed that the synchronized gamma phases in the IFJ were advanced by 20 ms compared to FFA and PPA, suggesting that the IFJ is the synchrony driver in a top-down manner. In an experiment using electrophysiology and pharmacological manipulations in monkeys, Bichot and colleagues showed that VPA plays a similar role in visual search to control non-spatial attention (Bichot et al., 2015). Simultaneous recordings from VPA, FEF and IT cortex revealed that the VPA shows the earliest time of feature selection. Therefore, they concluded that the VPA must be the source of the feature selection. Supporting parallel mechanisms for the control of spatial and non-spatial attention, pharmacological deactivation of the VPA impaired feature selection but, contrary to the effects of FEF inactivation, had no significant effect on spatial attention (Bichot et al., 2015). Taken together, in contrast to the oculomotor network-driven spatial attention signals, the above-mentioned studies demonstrated that attentional modulations seen in IT cortex only contribute to object recognition. Nevertheless, despite all of these recent advancements in our understanding of the operations in IT cortex, most of the previous studies have neglected its potential role beyond object recognition through fast feedforward converging processing (DiCarlo et al., 2012). To be more specific, the possibility that IT cortex, or more generally speaking the temporal cortex structures, is an active participant in attentional control beyond simply being a target of such control has rarely been considered.

#### 4. Temporal cortex participates in attentional control

One of the first studies showing that the IT cortex is involved in attentional selection was carried out by Rothblat and Pribram (1972). They trained monkeys on a task which demanded selecting either the color or the form of a complex stimulus. They showed that initially the neural responses from IT cortex time-locked to the response, and as it was not stimulus-locked, could play a role in stimulus selection. However, the occipital cortex also showed selective responses to the attended dimension which were instead locked to the stimulus onset. These results shed further light on the mechanism by which attention can lead to selection: early input filtering via the occipital cortex as a consequence of learning or later response filtering via the IT cortex (Rothblat and Pribram, 1972). It is this latter mechanism that suggests areas within IT would provide a basis for guiding attention and shaping behavior. Other studies add further support by showing that some neurons in aIT and pIT keep track of the behavioral relevance of visual features independent of their physical properties contributing to control of visual attention (Braitman, 1984; Fuster and Jervey, 1981). Attentional signals found in the temporal cortex were not limited to visual features such as color or form. Many cells in different regions of STS and IT including TEO were responding when attention had to be allocated to the fixation spot (Watanabe and Iwai, 1991). The same neurons were significantly less active when the fixation spot was blanking (Watanabe and Iwai, 1991). These studies provided the earliest evidence that necessary elements for the deployment of visual attention were likely present in IT cortex.

Neuroimaging studies have also highlighted a potential role for temporal cortex in controlling spatial and feature-based attention in both monkeys and humans (Corbetta et al., 1998; Caspari et al., 2015; Gitelman et al., 1999). As discussed in (Serenio et al., 2020), one reason that prior neurophysiological studies may have missed the contribution of the temporal cortex to attentional control is that they mainly focused on single neuron responses rather than how attention is reflected in the population responses. In an attempt to study attentional modulations

using population coding, Sereno and Lehigh utilized an experimental paradigm in which monkeys had to either pay attention to the location or shape of a stimulus (Serenio and Lehigh, 2018). While single neurons in ventral and dorsal stream areas (aIT and LIP) were significantly modulated by attention, multidimensional scaling analysis at the population level showed a significant attentional effect (better discriminability between locations and shapes) only in the aIT. Hence, the strength of attentional modulation of shape and location at a single-cell level in the temporal cortex might have been underestimated in the past (see (Serenio et al., 2020)). Nevertheless, it is not only the strength of modulation in IT cortex that is relevant, but whether attentional control signals originate from it as well.

Evidence for the latter comes from studies showing that ablating of cortical structures that include TEO impairs selective attention (De Weerd et al., 2003a, 2003b). Lesioning areas V4 or TEO had a significant effect on the monkeys' ability to filter out distractor information which interfered with the discrimination of targets within several feature domains. These attentional deficits were around two times larger when both V4 and TEO were simultaneously lesioned. The same group later showed that V4 and TEO are essential for successful spatial generalization tasks (De Weerd et al., 2003a, 2003b). Furthermore, neural responses in more anterior regions of IT cortex are significantly altered after V4/TEO lesions (Bertini et al., 2004; Buffalo et al., 2005). So, in order to efficiently filter out a distractor which is simultaneously presented together with a target in the receptive of IT neurons requires attentional filtering occurring within earlier areas with smaller receptive fields, in this case, V4 and TEO. As per some of the above-mentioned theoretical models of visual attention, this filtering would incorporate both local circuitry and feedback from later stages of the ventral visual stream (Bundesen et al., 2005; Tsotsos, 2011). We propose that this feedback originates in higher order areas and backpropagates to earlier visual areas.

#### 5. Temporal cortex attention network (TAN)

This section reviews recent evidence from behavioral, neurophysiological and fMRI studies that revealed distinct regions in the IT cortex and the STS, collectively referred to as the temporal cortex attention network (TAN), that guide visual attention in the presence of cognitive control signals.

##### 5.1. Area pITd

In a recent monkey fMRI study, area pITd, an area in the dorsal part of the posterior IT cortex (see Fig. 1), showed strong modulation during an attentive motion processing task (Stemann and Freiwald, 2016). This surprising finding contrasted with the general wisdom that attentional modulation of a ventral stream area is a consequence of drawing attention to the feature that area primarily processes (Moran and Desimone, 1985; Caspari et al., 2015; Patel et al., 2015), because area pITd had been shown to be sensitive to shape (Hikosaka et al., 1988) and color (Conway and Tsao, 2009), not motion. Despite the fact that motion was the main feature to be attended, motion-sensitive areas were much less modulated in their experiment. In a follow-up electrophysiological experiment (Stemann and Freiwald, 2019) they showed that area pITd exhibits the properties of a priority map encoding the spotlight of attention similar to areas LIP and FEF (Moore et al., 2003; Colby and Goldberg, 1999). They determined that: (1) three different tasks (motion detection, motion discrimination, and color discrimination) yielded similar attentional patterns. (2) Neurons within the pITd were not tuned to task-relevant stimulus features, like motion direction or color. (3) The activity of the pITd neurons was highly correlated with upcoming errors of attentional selection. (4) Microstimulation of pITd deployed spatial attention. This frames pITd as the first non-oculomotor area that exhibits the properties of a priority map, as described by Fecteau and Munoz (2006). Importantly, area pITd is a *feature-blind* priority map in the

ventral visual stream that can guide spatial attention. It is plausible that area pITd combines input from lower order feature conspicuity maps (such as area MT) and higher order areas involved in generation of top-down relevance signals to guide behavior.

A key factor of an attentional priority map is the ability to focus attention to specific locations within as well as across hemifields. As the recent pITd study (Stemann and Freiwald, 2019) only contained one target per hemifield, further studies need to show that the area pITd can deploy spatial attention to specific locations within the same hemifield (see (Moore and Fallah, 2001)) to fully establish pITd as a new priority map.

## 5.2. The STS area

Recent studies have delineated two distinct areas in the STS (see Fig. 1) which are implicated in the control of visual attention: (1) an area in the posterior superior temporal sulcus (pSTS), has been designated as the gaze following patch (GFP) as it has been demonstrated to have a key role in controlling spatial attention during social interactions in humans and monkeys (Ramezanzpour and Thier, 2020; Marciniak et al., 2014; Marquardt et al., 2017; Ramezanzpour et al., 2021) and (2) the middle STS region (mid-STS) in monkeys which has been shown to control spatial attention during motion direction and orientation discrimination tasks (Bogadhi et al., 2018, 2019).

### 1 GFP

The GFP located in the pSTS has been shown to specifically support the faculty of joint attention (Marciniak et al., 2014; Marquardt et al., 2017; Laube et al., 2011), a necessary component of primates' social development (Emery et al., 1997; Klein et al., 2009; Leopold and Krauzlis, 2020). Ramezanzpour and Thier performed the first electrophysiological recordings from the monkey GFP showing that neurons in this area are spatially tuned to locations that someone else is looking at, enabling the observer to shift their attention to the same (Ramezanzpour and Thier, 2020; Ramezanzpour et al., 2021). Furthermore, the study implicated GFP activity in cognitive gaze following behavior, highlighting the role of social context in shaping pSTS neural activity. In a control condition, in which the same visual stimuli were presented to the monkeys but they had to ignore the gaze cues and instead use the portrait's identity information to find the correct target, the same neurons that had been found to be spatially selective for gazed-at locations lost their spatial tuning. This result highlights the role of the GFP in cognitive control of gaze following according to task demands (Ramezanzpour and Thier, 2020; Ramezanzpour et al., 2021). Furthermore, the GFP neurons could target more than one gaze location within each hemifield, a finding which has yet to be shown in area pITd.

These findings not only suggested that the GFP plays a key role in a circuit controlling spatial attention, but they could also show the importance of this area in contextual control of attention. In their experimental paradigm, there was a period in which the animal had to pay attention to the color of a small dot. Different colors implied different cognitive strategies the animal had to take to solve the task, either following the subsequent face's gaze or to map his identity onto the same spatial targets layout which differed with the gazed-at location in most of the trials. The population of neural responses in the GFP was distinctively sensitive to the two instruction colors, red and green, and highlighted the importance of this area in the control of attention based on non-spatial contextual information. Notably, the differential responses to the color of the central fixation were tightly linked to behavioral performance as the animals made significantly more errors when the two separate populations of instruction selective neurons lost their selectivity. These findings show that the GFP is involved in determining the relevance of the social cues such as faces and hands to behavior in nonverbal communications (Ong et al., 2020) or predictability of social interactions (Roumazeilles et al., 2021).

### 2 mid-STS

Further evidence for the involvement of non-oculomotor structures in the control of spatial attention comes from two recent studies investigating consequences of midbrain SC inactivation on the rest of the brain activity during two covert attention tasks (Bogadhi et al., 2019, 2020). Previous studies of the same group showed significant attention-related modulations in monkeys' middle parts of the STS (mid-STS), a cortical area not traditionally linked to attention, during an attentive motion discrimination task (Bogadhi et al., 2018). The attention task required detection of a motion direction change in one of the two peripheral random motion stimuli. The color of the central cue was indicative of which stimuli should be covertly attended, and the monkeys had to report the relevant changes detected by releasing a joystick (Bogadhi et al., 2018). In a follow-up study, they showed that responses in the mid-STS and not motion-sensitive areas in the dorsal stream were most strongly attenuated after SC inactivation during the same tasks (Bogadhi et al., 2019). This reduction in the attentional modulation was replicated using another task replacing the random dot motion stimuli with second-order orientation stimuli that could not be discriminated according to the changes in motion energy. Finally, they inactivated the mid-STS region directly and observed attentional performance was similarly disrupted. Interestingly, in a follow-up study, the same group showed that many neurons in this area exhibit object selectivity and SC inactivation reduces this selectivity (Bogadhi et al., 2020). Altogether, these studies suggest that the mid-STS is involved in the SC's control of spatial attention, and thus is part of the hitherto oculomotor network's control of spatial attention. Interestingly, the mid-STS in monkeys has a functional connectivity fingerprint very similar to the temporoparietal junction (TPJ) in humans (Mars et al., 2013). Hence, its functionality might be reminiscent of the human TPJ i.e. reorienting attention towards a novel object of interest (Corbetta et al., 2008). However, the exact correspondence has yet to be confirmed. Nevertheless, the mid-STS attentional modulations are partially necessary for normal object representation in the temporal cortex (Bogadhi et al., 2020). It is important to note that the mid-STS area was located a few millimeters anterior to the reported anatomical coordinates of the GFP and pITd areas (Ramezanzpour and Thier, 2020; Stemann and Freiwald, 2019). The question of whether all of these three areas belong to the same cytoarchitectonic structure remains open.

Inactivation of the mid-STS caused monkeys to exhibit "spatial neglect", ignoring stimuli in the contralateral visual field (Bogadhi et al., 2019). While the early studies on spatial neglect suggested it was due to lesions to the temporo-parieto-occipital junction and inferior parietal lobule, later studies refined the location to the superior temporal cortex in which lesions cause spatial neglect (Karnath, 2001, 2015; Karnath et al., 2001). Thus, the neurophysiological inactivation of mid-STS results in monkeys are consistent with spatial neglect symptoms in humans as a consequence of broader STS lesions.

## 6. What do the TAN regions have in common?

One interesting common denominator of the above mentioned studies on the TAN is that all of the behavioral tasks have an executive control component i.e. a gaze should be followed or not (Ramezanzpour and Thier, 2020), a lever should be released or not (Bogadhi et al., 2019), a saccade should be made or not (Stemann and Freiwald, 2019), and all of these functions must be synchronized with temporal requirements of the tasks to ensure flexible performance. When the cognitive control signal interacts with attentional control, the TAN is engaged to support flexible behavior, a finding very similar to what has been already found in area LIP (Oristaglio et al., 2006; Gottlieb, 2012). Previously Oristaglio and colleagues had found some neurons in area LIP integrate covert spatial attention and a learnt stimulus-action association (Oristaglio et al., 2006). In their task, the animals had to release a bar held in their right paw if the cue was oriented to the right or a bar

held in the left paw if it was oriented to the left. A considerable number of LIP neurons responded to the attended location, something typical for LIP neurons, and the bar release. As discussed in depth in (Gottlieb, 2012), these neurons might be the basis of target selection interfaces with higher order processes of executive control which facilitate relevance assignment to visual cues. Indeed, the TAN neurons might have inherited their mixed selectivity (spatial attention and executive control) properties from their tandem LIP neurons or vice versa. It should be noted that such integrations are likely dependent on the visual working memory signals, which can contain task-relevant instructions and progression of the representations from the earlier visual areas up to the higher-level areas (Fuster, 1990). Indeed, a study revealed that successfully retaining visual information in working memory depends on a corticocortical loop of the prefrontal and IT cortex since bilateral cooling of each of these areas induced, in the other region, changes of spontaneous and task-related neuronal responses which were accompanied by lower performance in a working memory task (Fuster et al., 1985). This bidirectional interaction between the temporal cortex and the prefrontal fits well some of the theoretical models of attention in which bottom-up signals from lower-level visual areas and top-down signals from higher-level areas are both needed to select a target (Bundesen et al., 2005; Schneider, 2013; Tsotsos, 2011).

Another common denominator of the above-alluded studies on the GFP and pITd is that they need to deal with spatial transformations to generate a saccade (Sajad et al., 2020). By a closer look at the behavioral tasks used in the studies on the TAN, it becomes clear that output spatial coordinates derived from visual processing at the focus of attention (often in the periphery) need to be integrated with extra-retinal signals, such as the current eye position, to enable generating a precise saccadic eye movement initiating from the fixation point. Such spatial transformations were first observed at the level of the parietal cortex via gain modulation (Andersen and Mountcastle, 1983; Andersen et al., 1985). This concept, called “gain fields”, was later proposed to be important for invariant object recognition with the modulatory quantity (gain) being attention and led to the idea that IT cortex may use an attention-centered rather than eye-centered mechanism for invariance in object recognition (Pouget and Sejnowski, 1997; Salinas and Abbott, 1997; Salinas and Thier, 2000). An attention-centered coordinate system encodes location of objects relative to the current focus of attention when it does not align with the current gaze location (eye-centered coordinate). Gain modulation provides a basis for the idea that fewer receptive fields that can move around (attention-centered) independent of eye position can be an alternative solution for having too many receptive fields that are fixed at specific retinotopic locations (Salinas and Abbott, 1997; Salinas and Thier, 2000). As already mentioned in Section 1, a theoretical model developed by Bundesen and colleagues also proposes that dynamic remapping of receptive fields of cortical neurons is the basis for setting the weights which are used for reallocation of attention (Bundesen et al., 2005; Bundesen, 1990).

We predict that the TAN might contribute to such gain field transformations in a more general way as they can deal with spatial processing in parallel to complex features analysis as opposed to parietal neurons which are not involved in object recognition. The importance of spatial transformations, especially in social interactions, has been discussed in depth by Chang (2013).

Executive control of visual attention must deal with fixations (overtly or covertly), integration of bottom-up with top-down information (such as rules or prediction signals), spatial localization, priming, and other ingredients of visual tasks while enabling precise timing of the overall behavior (Tsotsos et al., 2021). “Cognitive programs”, or executive controllers, have been proposed to provide mechanistic integration of the above algorithms to ensure flexible behavior. The tasks recruited in studies of temporal cortex control of attention (Bogadhi et al., 2018, 2019, 2020; Ramezani and Thier, 2020; Stemmann and Freiwald, 2016, 2019; Marquardt et al., 2017) have several attentional elements such as cueing, priming, selection, covert or overt attention, endogenous

or exogenous initiation, spatial localization, disengaging and shifting attention, and surround suppression (for a full list of attentional elements see (Tsotsos et al., 2021; Tsotsos and Kruijne, 2014)). Hence it is necessary that a cognitive program oversees the whole process. The cognitive programs concept, proposed by Tsotsos and Kruijne (2014), is an advanced version of Ullman’s visual routines (Ullman, 1984) and emphasizes the fact that attention is much more complex than just selecting a region of interest for gaze fixation. Hence, active vision requires a set of algorithms beyond simple relationships between extracting shapes and spatial relationships to ensure it reaches its goal. The recently discovered TAN might indeed be representative cognitive programs for these types of tasks or at least contribute substantially though which exact aspects of cognitive programs are embedded in those areas must be determined through future studies. It is particularly important to investigate whether they are innate or they are developed via learning, where they are stored, and how they are retrieved. As discussed in Section 1, the theory of visual attention (Bundesen et al., 2005; Bundesen, 1990) assumes that there two successive neural representation of objects, a first one which is not accessible for goal directed action (proto-objects) and a second visual working memory representation (visual tokens) (Schneider, 2013). It is plausible to speculate that the TAN might indeed be corresponding to visual tokens, preparing the object to be acted on. This notion gets further support from the fact that neurons in the pITd were not representing visual motion or many of the GFP neurons were not representing faces (Ramezani and Thier, 2020; Stemmann and Freiwald, 2019). It would be interesting to test neurons in the TAN in working memory tasks more systematically in future studies. Relatedly, the prefrontal-like responses of the TAN (such as rule selectivity and context dependency of the GFP neurons) may suggest that the TAN may correspond to option identifier in hierarchical reinforcement learning models that route option-specific policies corresponding to stimulus-response pathways (Botvinick et al., 2009). Hierarchical reinforcement learning extends the classical models by assuming the learning agent’s actions can be made of reusable sub-routines or skills (Botvinick et al., 2009).

## 7. The TAN is distinct from dorsal and ventral attention networks

Previously, alternative models for visual attention control in humans, such as dorsal and ventral attention networks (DAN and VAN (Corbetta and Shulman, 2002),) attempted to explain where signals associated with top-down cognitive control might be integrated with bottom-up feature salience. Corbetta and Shulman suggested that while top-down signals for control of visual attention are generated in the dorsal posterior parietal and frontal regions (DAN), the VAN, including temporoparietal junction (TPJ) and inferior frontal cortex, directs attention to salient events in a more bottom-up manner and normal vision requires both systems to interact (Corbetta and Shulman, 2002). A follow-up fMRI study did not find functional evidence of a TPJ in macaque monkeys when testing them using the same paradigms which activated the human TPJ (Patel et al., 2015). Interestingly, another study revealed that the mid-STS area in macaques, part of the TAN, has the same functional connectivity profile as the human TPJ (Mars et al., 2013). These observations may raise the question of whether the monkey TAN might be indeed homologous to the posterior part of the human VAN i.e., the TPJ. We think this correspondence does not fit the functional properties of the monkey TAN and the human VAN for two main reasons: (1) The human VAN has been shown to be driven mainly by low-level features of the stimulus, such as color, in a bottom-up manner and regardless of the ongoing task, e.g., when they appear outside the cued focus of spatial attention unexpectedly (Corbetta and Shulman, 2002). However, the monkey TAN has been shown to be driven by more complex features, such as other’s gaze (Ramezani and Thier, 2020), and cognitive processes at the focus of the attention in a task-dependent manner, such as paying attention covertly to the movement direction of

stimulus while ignoring another stimulus in the opposite visual field over a long time (Bogadhi et al., 2019; Stemmann and Freiwald, 2019). (2) Neuroimaging studies have more consistently reported right-hemispheric activation of the VAN (see also for a review (Vossel et al., 2014)), however, the monkey TAN is bilaterally activated. When it comes to comparing attentional networks between monkeys and humans, it is noteworthy that the expansion of the temporal cortex in humans during the course of evolution in order to accommodate language, more complex social interactions, and some attention capacities not required for monkeys might have caused the above-mentioned asymmetrical VAN, its distinct functionality (right hemisphere dominance), and shifting the TAN areas farther away from each other in humans i.e. the monkey GFP and pITd are much closer together than the human GFP and phPIT (see Fig. 2).

Of course, the next logical step to investigate if there might be a human TAN, neither overlapping functionally nor anatomically with the human VAN, would be to conduct comparative fMRI studies in humans, now using the same behavioral paradigms as in monkey studies on the TAN. With the same logic, two studies attempted to localize the TAN regions in the human cortex (Marquardt et al., 2017; Sani et al., 2021). These two studies revealed that the areas corresponding most closely to the monkey GFP and pITd are clearly far from the human TPJ. The human GFP and the putative human posterior inferotemporal area (phPIT), constituting the human TAN, are located much more inferior to the human TPJ (see Fig. 2). Altogether, these observations confirm that there exists a TAN in both monkeys and humans, functionally and anatomically segregated from the posterior member of the VAN, i.e. the TPJ. Notwithstanding these similarities, future studies investigating connectivity and the causal role of the human TAN in control of attention are needed to ultimately conclude whether the monkey TAN and the human TAN are indeed homologous.

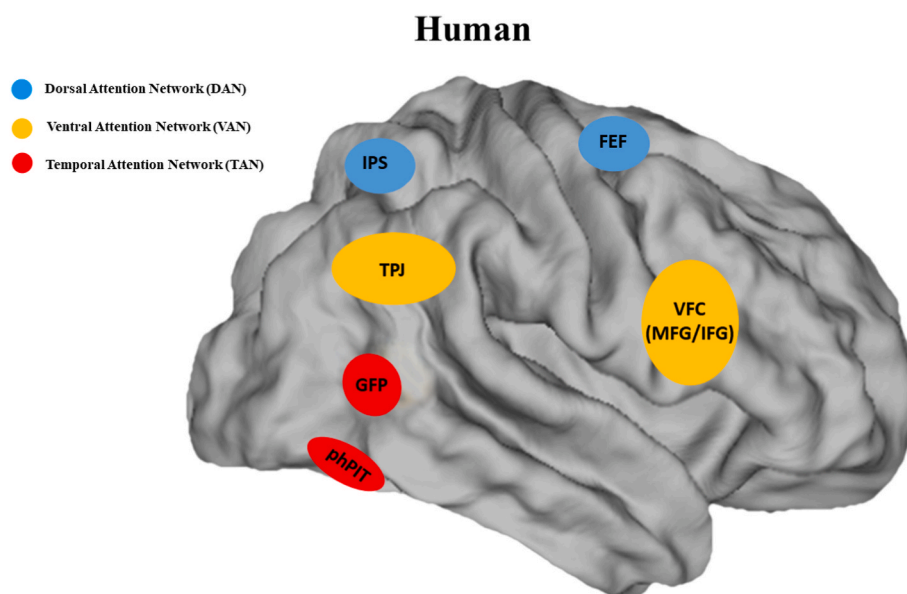
While the functional segregations of attention networks (DAN, VAN, and TAN) could explain some of the empirical observations, the mechanism by which these putative networks should interact is not well understood. The “Coherence field” notion has been able to provide a plausible solution (Serences and Yantis, 2006). This concept proposes that selective attention synchronizes the activity of neurons across topographically organized stages of the hierarchy to form coherence fields in which different areas contribute complementary information to support target selection. The coherence fields also support the notion that the relative influence of bottom-up stimulus features and top-down behavioral goals on the concerted activity of neurons in the visual

hierarchy varies along a continuum. Therefore, finding an exact locus in the hierarchy at which a shift from “source” to “target” occurs is complex (Serences and Yantis, 2006). This idea fits the selective tuning model (Tsotsos et al., 1995), which implies that the focus of attention is present throughout the whole neural network, starting at V1 and ending at the frontal cortex but with different resolutions (spatial and feature-based) and recurrent processing.

## 8. What could explain the functional properties of the TAN?

### 8.1. Feature specialized circuits

The anatomical location of the strong attentional signals found in the previously described recent studies of the mid-STS, GFP, and pITd is not identical, albeit very close (Bogadhi et al., 2019; Ramezani and Thier, 2020; Stemmann and Freiwald, 2019). What mechanisms could explain the strong attentional control signals found in these studies and not in more anterior parts of the temporal cortex such as aIT? The first important element is the degree of these areas’ (or their immediate neighbors) specialization with respect to different behavioral tasks employed. Emergence of highly specialized circuits to optimize neural information processing has been shown to be tightly linked to brain topology (Bullmore and Sporns, 2012) and the temporal cortex is not an exception. Current hypotheses suggest that at least the GFP is a domain-specific module since it controls spatial attention only based on social gaze information (see (Ramezani and Thier, 2020) for discussions). Whereas the pITd or mid-STS areas have been shown to be more generic as they respond to a wider range of attention tasks (such as attentive motion and color discrimination). While the full characteristics of each area specialization needs to be explored in more detail, importantly they are all adjacent to areas specialized with respect to the feature of the task being performed. For example GFP, with its dependence on social gaze, is in between posterior and middle face patches (Marciniak et al., 2014). The same holds for areas pITd and mid-STS which are sitting next to motion (MT, MST and FST) processing areas (Bogadhi et al., 2018; Stemmann and Freiwald, 2016). This close proximity to feature conspicuity maps may empower the attention control areas to access and integrate the necessary streams of information via local circuitries. Similarly, one can speculate that paying attention to highly complex shapes such as a face identity may recruit more anterior regions of the IT cortex specialized in facial identity processing. As a matter of fact, an fMRI study has already shown that



**Fig. 2.** Cortical attention networks in humans. Blue: The dorsal attention network (DAN) (Corbetta and Shulman, 2002; Corbetta et al., 2008; Vossel et al., 2014), red: the temporal attention network (TAN) (Marquardt et al., 2017; Sani et al., 2021), and orange: the ventral attention network (VAN) (Corbetta and Shulman, 2002; Corbetta et al., 2008; Vossel et al., 2014). IPS: intraparietal sulcus; FEF: frontal eye field; GFP: gaze following patch; phPIT: putative human posterior inferotemporal area; TPJ: temporoparietal junction; VFC: ventral frontal cortex; MFG: middle frontal gyrus; IFG: inferior frontal gyrus. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

when monkeys were instructed to use other monkeys' facial identities to pinpoint particular spatial targets associated with those identities, an area in the anterior temporal cortex, likely one of the anterior face processing patches, was activated (Marciniak et al., 2014). When considering that idea that the location of the TAN regions may depend on the adjacent conspicuity maps, to take advantage of the local information processing mechanisms already available in those areas (such as lateral inhibition, and etc.), one can speculate the TAN can also serve auditory attention as it is also close to auditory cortical areas (Petkov et al., 2006). Previous studies have already shown that auditory attention operates similar to visual attention (Kayser et al., 2005). Indeed, the TAN could be a site of convergence for deploying attention to integrated auditory and visual information and thus further facilitating social interaction processing.

Still, one question remains for future studies: Are the attention signals observed in the temporal cortex driving spatial attention (location) or object-based attention (the object itself)? This question can be addressed by dissociating spatial and object-based attention in a scenario such as priming spatial attention in the absence of visual information (see (Kastner et al., 1999) as an example).

We should emphasize that while we presented the findings of the recent studies as three separate areas, there is converging evidence that specific cognitive operations are emergent property of network operations rather than being strictly linked to activity in restricted parts of the brain. Future studies are required to investigate how these areas interact with the rest of the brain, including the VAN and DAN, during cognitive operations.

## 8.2. Visual field maps

A visual field map is a representation that might be used for control mechanisms and could explain the distinct loci of attentional control signals beyond different behavioral tasks (and different features). Early imaging studies on the retinotopic organization of the visual cortex did not consistently find retinotopy in IT cortex. The initial lack of retinotopic maps in other parts of the brain other than early visual areas could have been either due to the low signal to noise ratio of the neuroimaging technology or not using an appropriate protocol which considers stimulus-based selectivity in addition to the visual field location (Saygin and Sereno, 2008). However, this notion was later refined by some studies showing a systematic representation of eccentricity in the ventral stream (Levy et al., 2001; Arcaro and Livingstone, 2017; Hasson et al., 2002, 2003; Janssens et al., 2014; Kolster et al., 2014). To date, the existence of retinotopic maps have been shown in parietal, and even frontal cortex (Mackey et al., 2017; Wang et al., 2015b). These findings are important as one can hypothesize that retinotopy might constitute a basis for connecting different parts of the frontal, parietal and temporal cortex dealing with the same part of the visual field together.

Note that in each of the above studies attention had to be deployed to a certain location in the visual field. In the studies on the pITd the two stimuli were presented at five degree eccentricity (Stemmann and Freiwald, 2016, 2019), while in the studies on the mid-STS the stimuli were presented at eight degree eccentricity (Bogadhi et al., 2018, 2019), and the GFP study included a central stimulus which had to be attended foveally (Ramezani and Thier, 2020). Apart from other behavioral demands which varied across these studies, the eccentricity perhaps had played a major role in the location of attention-related activities in the temporal cortex. The fact that the GFP area was very close to the posterior part of the temporal cortex which has a foveal bias could be because they presented the face stimuli at the center of the visual field (see Fig. 1). Similarly, the pITd and mid-STS studies used more eccentric stimuli, which could have caused a shift in the locus of the attentional modulation according to the visual field representation and eccentricity gradient in the posterior temporal cortex towards more anterior and ventral parts of the temporal cortex (see (Conway, 2018) for a review). If the stimulus eccentricity drives the cortical activation pattern, then one

can predict that using more eccentric locations for stimulus presentation may shift the locus of attentional control even further anterior in the temporal cortex. Future experiments are needed to test this hypothesis.

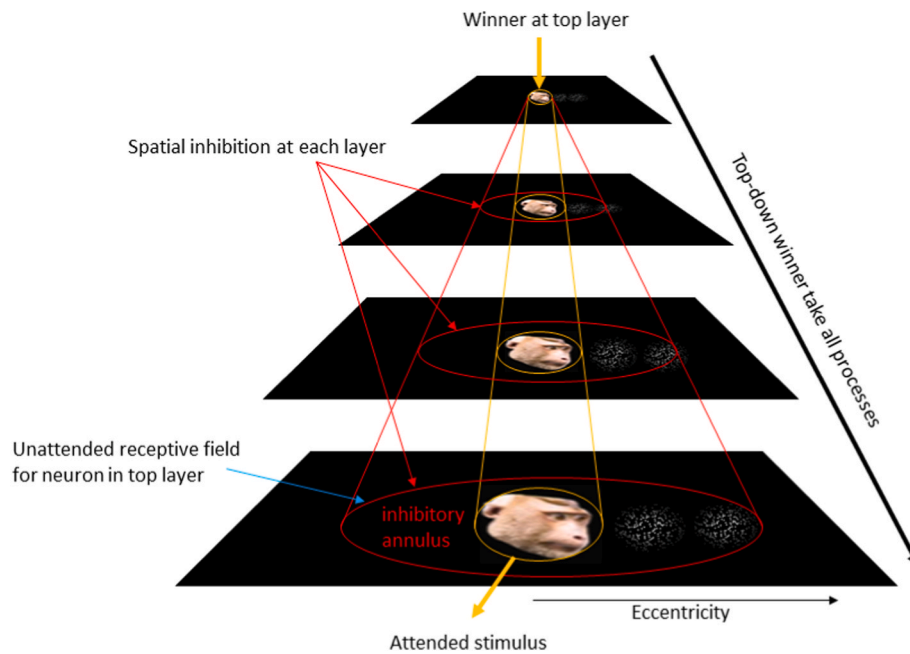
As described in the selective tuning model, spatial attention is distributed both within a given layer (area within the visual processing hierarchy) as well as feeding back from higher level areas. The existence of retinotopic maps in the frontal and parietal cortices (Mackey et al., 2017) might suggest that areas with the same eccentricity bias in prefrontal and parietal cortices may drive visuospatial attention in the posterior-middle temporal areas which in turn modulate the lower level visual areas even as early as V1 in a top-down manner. In this framework, the temporal cortex attentional signals can be considered as intermediate feedback signals necessary for the formation of high spatial resolution focus of attention in earlier retinotopic areas of the visual hierarchy. In fact, the distinct topography of the attentional control areas seen in the posterior-middle temporal cortex could also be a consequence of the surround suppression as described in the selective tuning model (Tsotsos, 2011) or the normalization model of attention (Reynolds and Heeger, 2009). Each stimulus is presented and attended at a certain eccentricity which suppresses the immediate part of visual field around it in a retinotopic manner. Hence the neighboring areas in the retinotopic map are suppressed (Fig. 3). That would explain GFP suppression of the neighboring areas pITd and mid-STS in fMRI activation maps. However, future studies are needed to perform all of these tasks and retinotopic mapping on the same set of subjects to confirm whether the GFP, pITd, and mid-STS areas together form a complete map of the visual field.

## 9. Connectivity of the monkey TAN

The cortical and subcortical connectome of the TAN, specifically to the two other classical attention networks (DAN and VAN) has yet to be fully established. Nevertheless, there is already some evidence which might help to better understanding the TAN's wiring to the rest of the brain. In monkeys, pITd's connectivity further supports its ability to be part of the network controlling spatial attention. Diffusion tensor imaging (DTI) showed that white matter bundles originating in pITd are well connected to the two nodes of the DAN: FEF and LIP (Sani et al., 2019). Tract tracing studies have shown that the anatomically defined area pITd has feedback projections to earlier visual areas such as V4, V3, V2 and even V1 via either V4-V3 or V4-V2 (Distler et al., 1993). Connectivity of the functionally defined GFP has not yet been studied. Nevertheless, purely anatomical studies have shown that in addition to reciprocal connections to the early visual areas, the GFP receives input from a subcortical pathway including the pulvinar (Kaas and Lyon, 2007) and SC (Bogadhi et al., 2019, 2020) and in turn projects to the interparietal sulcus, likely area LIP (Baizer et al., 1991). Cortex around the mid-STS has been shown to connect visual areas such as V2, V3, MT, and MST, to more anterior parts of the STS (Boussaoud et al., 1990). The mid-STS is also connected with attention-related areas such as FEF (Boussaoud et al., 1990) and the pulvinar (Kagan et al., 2021). Hence, one can conclude that the mid-STS is the central node of a pathway connecting cortical motion processing (starting from MT) with subcortical and cortical attentional control areas.

While anatomical studies using tract-tracing are informative, they largely depend on the coarse parcellation of cytoarchitectonic areas. Moreover, it is often difficult to conclude whether an anatomically defined area may exactly fit to its fMRI-delineated functional correspondence. One method which can overcome these limitations is combined electrical stimulation-fMRI (Tolias et al., 2005; Xu et al., 2021). While concurrent electrical stimulation-fMRI of the functionally defined TAN has yet not been performed, a recent study focusing on the lateral prefrontal cortex (LPFC) revealed that there exists a topographic progression of the LPFC connections on the cortical surface along a particular direction (Xu et al., 2021) such that a caudoventral to rostroventral gradient in temporal cortex and a caudal/sulcal to





**Fig. 3.** Attentional surround suppression. Attending to a face presented foveally activates the GFP and generates a suppressive surround for stimuli presented at larger eccentricities and deactivates the neighboring pITd and mid-STs areas.

rostral/superficial gradient in posterior parietal cortex seem to map to a largely overlapping topographical map in the LPFC. Based on this map, the GFP and the mid-STs are likely connected to FEF and its anterior/ventral neighbor. The non-retinotopic continuum in the LPFC which integrates retinotopic maps of the temporal and parietal cortices, can fulfill the requirements of the top layer in various theoretical models to establish the flow of top-down signals for guiding attention (Bundesen et al., 2005; Tsotsos et al., 2021; Miller and Buschman, 2013). The TAN as an intermediate level of the attentional control hierarchy, can continuously link early representations in the visual cortex and subcortical areas which are more robust with higher level flexible cognitive processes in the parietal and prefrontal cortices to ensure an optimal behavior.

## 10. Implications of active vision in temporal cortex for social interactions

Intelligent social behavior requires flexibly attending to cues from multiple sensory modalities provided by the other individual. For instance, in one moment we may assess the other's gaze direction in order to identify his/her focus of attention, whereas, in the next moment we may focus on his/her hand pointing toward a certain object being discussed. There is converging evidence that the primate brain treats social information differently to the extent that the existence of a third visual pathway has been recently hypothesized (Pitcher and Ungerleider, 2021). This third visual pathway which connects the early visual areas to the STS, plays a crucial role in processing dynamic aspects of social perception such as moving faces and bodies, ultimately leading to the understanding of others' actions and theory of mind (Wyk et al., 2009).

We hypothesize that the participation of the TAN regions in attentional control and cognitive programs is also beneficial for action understanding. In order to arrive at a complete interpretation of the given social context, we need to flexibly switch between various cues and actions and integrate the information collected. Hitherto, the neurophysiological principles that orchestrate ensembles of neurons to flexibly link these cues to generate a meaningful and dynamic percept of other's actions have been mostly studied in the context of mirror neurons (Rizzolatti and Craighero, 2004). Nevertheless, there are several

studies showing that various regions in the temporal cortex contribute to the perception of other's actions in both monkeys and humans (Iacoboni et al., 2001; Perani et al., 2001; Pierno et al., 2006; Kilintari et al., 2014; Nelissen et al., 2006, 2011), similar to what has previously been found in the premotor cortex (Rizzolatti and Craighero, 2004). Importantly, some of these areas even show activity during execution of the same actions (Kilintari et al., 2014). While some of these areas such as the STP, which encompasses regions TPO, PGa and IPa, are active during action observation and execution only when actions are visible (Kilintari et al., 2014), some other areas such as MT, MST, FST, remain active even if actions are performed in darkness or if they are invisible to the performer (Kilintari et al., 2014; Gazzola and Keysers, 2009). These findings suggest that mirror-like responses found in the motion complex part of the temporal cortex (MT, MST, FST) may reflect visual imagery from the actor's point of view (Kilintari et al., 2011, 2014) which rely on top-down efferent signals from prefrontal and parietal cortices. Interestingly, at least the mid-STs, a member of the TAN, partially overlaps with FST and parts of the TPO.

In contrast to the temporal cortex, other brain regions such as pre-motor and posterior parietal cortex (Rizzolatti and Sinigaglia, 2010), orbitofrontal cortex (Azzi et al., 2012), anterior cingulate cortex (Chang et al., 2013; Yoshida et al., 2012; Haroush and Williams, 2015) have so far been considered to play a more important role when it comes to social action monitoring, action observation, and action execution. It remains to be investigated how the TAN and other areas of the STS which showed mirror-like responses functionally relate to these areas in the parietal and frontal cortices.

How can the TAN contribute to action understanding during social interactions? Primates as a social species need to process and direct attention based on social cues as much as nonsocial ones (such as a flashing red light). At a lower level, attention has been shown to play an important role in binding visual features such as color and motion into an object representation (Treisman and Gelade, 1980; Bodelón et al., 2007; Perry and Fallah, 2014). Hence it is parsimonious to assume that attention might also be necessary to bind different social cues provided by others in order to generate a unified action. Take gaze following as an example in which attention should be constantly paid to the other person's face in order to be able to detect abrupt changes in their gaze direction, head movements, and finally the object they fixate on. As

described previously, the processes which are needed to perform these actions rely on cognitive programs to which the temporal cortex contributes. Consistent with the selective tuning model of attention, temporal cortex, as an intermediate level in the visual hierarchy, can reduce interference among other elements by operations such as pruning away task irrelevant information. The idea that TAN could bind several representations of social cues into one unified action gets further support from a theory by Keyzers and Perrett that hypothesizes events which systematically follow each other could be associated in Hebbian ways across various modalities such as visual and motor domains. They further proposed that there are action codes which integrate visual effects with motor commands and it is possible that Hebbian learning retrieves the STS representation of actions during executions (Keyzers and Perrett, 2004).

Why is the temporal cortex a good candidate for implementing the above operations? The essential elements of social interactions such as faces, body parts, and biological motion are indeed already represented in the temporal cortex. Hence implementing attention control signals and cognitive programs at the level of temporal cortex is ecologically beneficial because the local information processing mechanisms already available in these areas (such as lateral inhibition, surround suppression, etc.) can be recruited for binding social cues, flexibly switching between them, and implementing them into attentional control.

## 11. Conclusions

Through reviewing the latest findings on the potential role of temporal cortex in guiding visual attention, we propose that the STS and IT cortex participate in target selection and cognitive programs. We reviewed how different behavioral tasks with different visual stimuli presented at different eccentricities may explain attentional control signals observed in the posterior and middle STS and IT cortex regions, collectively referred to them as the TAN.

However, as this perspective on active vision by temporal cortex generalizes the existence of attention controllers to non-oculomotor structures, new questions arise that cannot be addressed without further study. Perhaps the most important question is what differences, if any, exist between the temporal cortex control of attention and previously known oculomotor system attention control areas. Most of the previous studies which missed attentional control signals in the temporal cortex had used relatively simple tasks and focused on control areas that produce motor output since they are easier to compare to behavior. However, it is becoming more evident that complex object representations can guide motor systems which do not represent that complexity (see (Kehoe et al., 2021) as an example from the oculomotor system). This notion supports the existence of cognitive programs in temporal cortex which are necessary to work in conjunction with oculomotor priority maps in frontal and parietal cortices to produce complex behaviors. The use of cognitive programs is a parsimonious solution to the brain as a dynamical system which needs to deal with a wide range of complex tasks and behaviors. Different nodes of this system play different roles for specific tasks, while the connectivity ensures the relevant control signals reach their target(s). The area that is most able to differentiate the visual scene into different priorities would be the one to drive the overall system. This view also implies that the focus of attention is present throughout the whole neural network starting at V1 and ending at frontal cortex but with different resolutions (spatial and feature-based) and recurrent processing. Control of visual attention is not just a function of the fronto-parietal motor control networks. Temporal cortex also constitutes to cognitive programs that deploy visual attention to earlier stages of the ventral stream to support target selection and enable flexible complex behaviors.

## No data statement

This paper does not contain original data or analysis.

## CRedit authorship contribution statement

**Hamidreza Ramezanzpour:** Writing – original draft, Writing – review & editing. **Mazyar Fallah:** Supervision, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgment

H.R. was supported by a Vision: Science to Applications (VISTA) postdoctoral fellowship award. M.F. was supported by an NSERC Discovery Grant (RGPIN-2016-05296) and CIHR Project Grant (MOP-102482).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.crneur.2022.100038>.

## References

- Andersen, R.A., Mountcastle, V.B., 1983. The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *J. Neurosci.* 3, 532–548.
- Andersen, R.A., Essick, G.K., Siegel, R.M., 1985. Encoding of spatial location by posterior parietal neurons. *Science* 230, 456–458.
- Arcaro, M.J., Livingstone, M.S., 2017. Retinotopic organization of scene areas in macaque inferior temporal cortex. *J. Neurosci.* 37, 7373–7389.
- Armstrong, K.M., Chang, M.H., Moore, T., 2009. Selection and maintenance of spatial information by frontal eye field neurons. *J. Neurosci.* 29, 15621–15629.
- van Atteveldt, N., Formisano, E., Goebel, R., Blomert, L., 2004. Integration of letters and speech sounds in the human brain. *Neuron* 43, 271–282.
- Azzi, J.C.B., Sirigu, A., Duhamel, J.-R., 2012. Modulation of value representation by social context in the primate orbitofrontal cortex. *Proc. Natl. Acad. Sci. Unit. States Am.* 109, 2126–2131.
- Baizer, J.S., Ungerleider, L.G., Desimone, R., 1991. Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *J. Neurosci.* 11, 168–190.
- Baldauf, D., Desimone, R., 2014. Neural mechanisms of object-based attention. *Science* 344, 424–427.
- Bao, P., She, L., McGill, M., Tsao, D.Y., 2020. A map of object space in primate inferotemporal cortex. *Nature* 583, 103–108.
- Beauchamp, M.S., Lee, K.E., Argall, B.D., Martin, A., 2004. Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron* 41, 809–823.
- Bertini, G., Buffalo, E.A., De Weerd, P., Desimone, R., Ungerleider, L.G., 2004. Visual responses to targets and distracters by inferior temporal neurons after lesions of extrastriate areas V4 and TEO. *Neuroreport* 15, 1611–1615.
- Bichot, N.P., Schall, J.D., 1999. Effects of similarity and history on neural mechanisms of visual selection. *Nat. Neurosci.* 2, 549–554.
- Bichot, N.P., Schall, J.D., 2002. Priming in macaque frontal cortex during popout visual search: feature-based facilitation and location-based inhibition of return. *J. Neurosci.* 22, 4675–4685.
- Bichot, N.P., Heard, M.T., DeGennaro, E.M., Desimone, R., 2015. A source for feature-based attention in the prefrontal cortex. *Neuron* 88, 832–844.
- Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S.F., Springer, J.A., Kaufman, J.N., Possing, E.T., 2000. Human temporal lobe activation by speech and nonspeech sounds. *Cerebr. Cortex* 10, 512–528.
- Bisley, J.W., Goldberg, M.E., 2003. Neuronal activity in the lateral intraparietal area and spatial attention. *Science* 299, 81–86.
- Bisley, J.W., Goldberg, M.E., 2010. Attention, intention, and priority in the parietal lobe. *Annu. Rev. Neurosci.* 33, 1–21.
- Bodelón, C., Fallah, M., Reynolds, J.H., 2007. Temporal resolution for the perception of features and conjunctions. *J. Neurosci.* 27, 725–730.
- Boehler, C.N., Tsotsos, J.K., Schoenfeld, M.A., Heinze, H.-J., Hopf, J.-M., 2009. The center-surround profile of the focus of attention arises from recurrent processing in visual cortex. *Cerebr. Cortex* 19, 982–991.
- Bogadhi, A.R., Bollimunta, A., Leopold, D.A., Krauzlis, R.J., 2018. Brain regions modulated during covert visual attention in the macaque. *Sci. Rep.* 8, 15237.
- Bogadhi, A.R., Bollimunta, A., Leopold, D.A., Krauzlis, R.J., 2019. Spatial attention deficits are causally linked to an area in macaque temporal cortex. *Curr. Biol.* 29, 726–736.e4.
- Bogadhi, A.R., Katz, L.N., Bollimunta, A., Leopold, D.A., Krauzlis, R.J., 2020. Midbrain activity shapes high-level visual properties in the primate temporal cortex. *Neuron*.

- Botvinick, M.M., Niv, Y., Barto, A.G., 2009. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition* 113, 262–280.
- Boussaoud, D., Ungerleider, L.G., Desimone, R., 1990. Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *J. Comp. Neurol.* 296, 462–495.
- Braitman, D.J., 1984. Activity of neurons in monkey posterior temporal cortex during multidimensional visual discrimination tasks. *Brain Res.* 307, 17–28.
- Buffalo, E.A., Bertini, G., Ungerleider, L.G., Desimone, R., 2005. Impaired filtering of distracter stimuli by TE neurons following V4 and TEO lesions in macaques. *Cerebr. Cortex* 15, 141–151.
- Bullmore, E., Sporns, O., 2012. The economy of brain network organization. *Nat. Rev. Neurosci.* 13, 336–349.
- Bundesen, C., 1990. A theory of visual attention. *Psychol. Rev.* 97, 523–547.
- Bundesen, C., Habekost, T., Kyllingsbaek, S., 2005. A neural theory of visual attention: bridging cognition and neurophysiology. *Psychol. Rev.* 112, 291–328.
- Buschman, T.J., Miller, E.K., 2007. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 1860–1862.
- Caspari, N., Janssens, T., Mantini, D., Vandenberghe, R., Vanduffel, W., 2015. Covert shifts of spatial attention in the macaque monkey. *J. Neurosci.* 35, 7695–7714.
- Chang, S.W.C., 2013. Coordinate transformation approach to social interactions. *Front. Neurosci.* 7.
- Chang, S.W.C., Gariépy, J.-F., Platt, M.L., 2013. Neuronal reference frames for social decisions in primate frontal cortex. *Nat. Neurosci.* 16, 243–250.
- Chelazzi, L., Miller, E.K., Duncan, J., Desimone, R., 1993. A neural basis for visual search in inferior temporal cortex. *Nature* 363, 345–347.
- Cohen, M.R., Maunsell, J.H.R., 2011. Using neuronal populations to study the mechanisms underlying spatial and feature attention. *Neuron* 70, 1192–1204.
- Colby, C.L., Goldberg, M.E., 1999. Space and attention in parietal cortex. *Annu. Rev. Neurosci.* 22, 319–349.
- Connor, C.E., Preddie, D.C., Gallant, J.L., Van Essen, D.C., 1997. Spatial attention effects in macaque area V4. *J. Neurosci.* 17, 3201–3214.
- Conway, B.R., 2018. The organization and operation of inferior temporal cortex. *Annu. Rev. Vis. Sci.* 4, 381–402.
- Conway, B.R., Tsao, D.Y., 2009. Color-tuned neurons are spatially clustered according to color preference within alert macaque posterior inferior temporal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 106, 18034–18039.
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.
- Corbetta, M., Miezin, F.M., Dobmeyer, S., Shulman, G.L., Petersen, S.E., 1990. Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* 248, 1556–1559.
- Corbetta, M., Akbudak, E., Conturo, T.E., Snyder, A.Z., Ollinger, J.M., Drury, H.A., Linenweber, M.R., Petersen, S.E., Raichle, M.E., Van Essen, D.C., et al., 1998. A common network of functional areas for attention and eye movements. *Neuron* 21, 761–773.
- Corbetta, M., Patel, G., Shulman, G.L., 2008. The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58, 306–324.
- Dahl, C.D., Logothetis, N.K., Kayser, C., 2009. Spatial organization of multisensory responses in temporal association cortex. *J. Neurosci.* 29, 11924–11932.
- Deen, B., Koldewyn, K., Kanwisher, N., Saxe, R., 2015. Functional organization of social perception and cognition in the superior temporal sulcus. *Cerebr. Cortex* 25, 4596–4609.
- Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222.
- Desimone, R., Albright, T.D., Gross, C.G., Bruce, C., 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. *J. Neurosci.* 4, 2051–2062.
- DiCarlo, J.J., Zoccolan, D., Rust, N.C., 2012. How does the brain solve visual object recognition? *Neuron* 73, 415–434.
- Distler, C., Boussaoud, D., Desimone, R., Ungerleider, L.G., 1993. Cortical connections of inferior temporal area TEO in macaque monkeys. *J. Comp. Neurol.* 334, 125–150.
- Dobs, K., Schultz, J., Bühlhoff, I., Gardner, J.L., 2018. Task-dependent enhancement of facial expression and identity representations in human cortex. *Neuroimage* 172, 689–702.
- Downing, P.E., Jiang, Y., Shuman, M., Kanwisher, N., 2001. A cortical area selective for visual processing of the human body. *Science* 293, 2470–2473.
- Ekstrom, L.B., Roelfsema, P.R., Arsenault, J.T., Bonmassar, G., Vanduffel, W., 2008. Bottom-up dependent gating of frontal signals in early visual cortex. *Science* 321, 414–417.
- Emery, N.J., Lorincz, E.N., Perrett, D.I., Oram, M.W., Baker, C.I., 1997. Gaze following and joint attention in rhesus monkeys (*Macaca mulatta*). *J. Comp. Psychol.* 111, 286–293. Washington, DC : 1983.
- Fallah, M., Stoner, G.R., Reynolds, J.H., 2007. Stimulus-specific competitive selection in macaque extrastriate visual area V4. *Proc. Natl. Acad. Sci. Unit. States Am.* 104, 4165–4169.
- Fecteau, J.H., Munoz, D.P., 2006. Saliency, relevance, and firing: a priority map for target selection. *Trends Cognit. Sci.* 10, 382–390.
- Fiebelkorn, I.C., Kastner, S., 2019. A rhythmic theory of attention. *Trends Cognit. Sci.* 23, 87–101.
- Fisher, C., Freiwald, W.A., 2015. Contrasting specializations for facial motion within the macaque face-processing system. *Curr. Biol. : Cailiao Baohu* 25, 261–266.
- Freiwald, W.A., Tsao, D.Y., 2010. Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science (New York, NY)* 330, 845–851.
- Fujita, I., Tanaka, K., Ito, M., Cheng, K., 1992. Columns for visual features of objects in monkey inferotemporal cortex. *Nature* 360, 343–346.
- Fuster, J.M., 1990. Inferotemporal units in selective visual attention and short-term memory. *J. Neurophysiol.*
- Fuster, J.M., Bauer, R.H., Jervey, J.P., 1985. Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Res.* 330, 299–307.
- Fuster, J.M., Jervey, J.P., 1981. Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science* 212, 952–955.
- Gaillard, C., Ben Hadj Hassen, S., Di Bello, F., Bihan-Poudec, Y., VanRullen, R., Ben Hamed, S., 2020. Prefrontal attentional saccades explore space rhythmically. *Nat. Commun.* 11, 925.
- Gazzola, V., Keysers, C., 2009. The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cerebr. Cortex* 19, 1239–1255.
- Gitelman, D.R., Nobre, A.C., Parrish, T.B., LaBar, K.S., Kim, Y.-H., Meyer, J.R., Mesulam, M.-M., 1999. A large-scale distributed network for covert spatial attention: Further anatomical delineation based on stringent behavioural and cognitive controls. *Brain* 122, 1093–1106.
- Goldberg, M.E., Bisley, J., Powell, K.D., Gottlieb, J., Kusunoki, M., 2002. The role of the lateral intraparietal area of the monkey in the generation of saccades and visuospatial attention. *Ann. N. Y. Acad. Sci.* 956, 205–215.
- Goodale, M.A., Milner, A.D., 1992. Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25.
- Gottlieb, J., 2002. Parietal mechanisms of target representation. *Curr. Opin. Neurobiol.* 12, 134–140.
- Gottlieb, J., 2012. Attention, learning, and the value of information. *Neuron* 76, 281–295.
- Gottlieb, J.P., Kusunoki, M., Goldberg, M.E., 1998. The representation of visual salience in monkey parietal cortex. *Nature* 391, 481–484.
- Gross, C.G., 1994. How inferior temporal cortex became a visual area. *Cerebr. Cortex* 4, 455–469.
- Gross, C.G., Bender, D.B., Rocha-Miranda, C.E., 1969. Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science* 166, 1303–1306.
- Hanning, N.M., Szinte, M., Deubel, H., 2019. Visual attention is not limited to the oculomotor range. *Proc. Natl. Acad. Sci. U. S. A.* 116, 9665–9670.
- Harroush, K., Williams, Z.M., 2015. Neuronal prediction of opponent's behavior during cooperative social interchange in primates. *Cell* 160, 1233–1245.
- Hasson, U., Levy, I., Behrmann, M., Hendler, T., Malach, R., 2002. Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 34, 479–490.
- Hasson, U., Harel, M., Levy, I., Malach, R., 2003. Large-scale mirror-symmetry organization of human occipito-temporal object areas. *Neuron* 37, 1027–1041.
- Hein, G., Knight, R.T., 2008. Superior temporal sulcus—It's my area: or is it? *J. Cognit. Neurosci.* 20, 2125–2136.
- Hikosaka, K., Iwai, E., Saito, H., Tanaka, K., 1988. Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. *J. Neurophysiol.* 60, 1615–1637.
- Hoffman, E.A., Haxby, J.V., 2000. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nat. Neurosci.* 3, 80–84.
- Humphreys, G.W., Duncan, J., Treisman, A., Desimone, R., 1998. Visual attention mediated by biased competition in extrastriate visual cortex. *Phil. Trans. Roy. Soc. Lond. B Biol. Sci.* 353, 1245–1255.
- Iacoboni, M., Koski, L.M., Brass, M., Bekkering, H., Woods, R.P., Dubeau, M.-C., Mazziotta, J.C., Rizzolatti, G., 2001. Reafferent copies of imitated actions in the right superior temporal cortex. *Proc. Natl. Acad. Sci. Unit. States Am.* 98, 13995–13999.
- Ibos, G., Freedman, D.J., 2014. Dynamic integration of task-relevant visual features in posterior parietal cortex. *Neuron* 83, 1468–1480.
- Ibos, G., Freedman, D.J., 2016. Interaction between spatial and feature attention in posterior parietal cortex. *Neuron* 91, 931–943.
- Ignashchenkova, A., Dicke, P.W., Haarmeier, T., Thier, P., 2004. Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. *Nat. Neurosci.* 7, 56–64.
- Ipata, A.E., Gee, A.L., Goldberg, M.E., 2012. Feature attention evokes task-specific pattern selectivity in V4 neurons. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16778–16785.
- Isik, L., Koldewyn, K., Beeler, D., Kanwisher, N., 2017. Perceiving social interactions in the posterior superior temporal sulcus. *Proc. Natl. Acad. Sci. Unit. States Am.* 114, E9145–E9152.
- Itti, L., Koch, C., 2001. Computational modelling of visual attention. *Nat. Rev. Neurosci.* 2, 194–203.
- Itti, L., Koch, C., Niebur, E., 1998. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 20, 1254–1259.
- Janssens, T., Zhu, Q., Popivanov, I.D., Vanduffel, W., 2014. Probabilistic and single-subject retinotopic maps reveal the topographic organization of face patches in the macaque cortex. *J. Neurosci.* 34, 10156–10167.
- Kaas, J.H., Lyon, D.C., 2007. Pulvinar contributions to the dorsal and ventral streams of visual processing in primates. *Brain Res. Rev.* 55, 285–296.
- Kagan, I., Gibson, L., Spanou, E., Wilke, M., 2021. Effective connectivity and spatial selectivity-dependent fMRI changes elicited by microstimulation of pulvinar and LIP. *Neuroimage* 240, 118283.
- Karnath, H.-O., 2001. New insights into the functions of the superior temporal cortex. *Nat. Rev. Neurosci.* 2, 568–576.
- Karnath, H.-O., 2015. Spatial attention systems in spatial neglect. *Neuropsychologia* 75, 61–73.
- Karnath, H.-O., Ferber, S., Himmelbach, M., 2001. Spatial awareness is a function of the temporal not the posterior parietal lobe. *Nature* 411, 950–953.
- Kastner, S., De Weerd, P., Desimone, R., Ungerleider, L.G., 1998. Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science* 282, 108–111.

- Kastner, S., Pinsk, M.A., De Weerd, P., Desimone, R., Ungerleider, L.G., 1999. Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22, 751–761.
- Kayser, C., Petkov, C.I., Lippert, M., Logothetis, N.K., 2005. Mechanisms for allocating auditory attention: an auditory saliency map. *Curr. Biol.* 15, 1943–1947.
- Kehoe, D.H., Lewis, J., Fallah, M., 2021. Oculomotor target selection is mediated by complex objects. *J. Neurophysiol.*
- Keysers, C., Perrett, D.I., 2004. Demystifying social cognition: a Hebbian perspective. *Trends Cognit. Sci.* 8, 501–507.
- Kilintari, M., Raos, V., Savaki, H.E., 2011. Grasping in the dark activates early visual cortices. *Cerebr. Cortex* 21, 949–963.
- Kilintari, M., Raos, V., Savaki, H.E., 2014. Involvement of the superior temporal cortex in action execution and action observation. *J. Neurosci.* 34, 8999–9011.
- Klein, J.T., Shepherd, S.V., Platt, M.L., 2009. Social attention and the brain. *Curr. Biol.* 19, R958–R962.
- Koch, C., Ullman, S., 1985. Shifts in selective visual attention: towards the underlying neural circuitry. *Hum. Neurobiol.* 4, 219–227.
- Koida, K., Komatsu, H., 2007. Effects of task demands on the responses of color-selective neurons in the inferior temporal cortex. *Nat. Neurosci.* 10, 108–116.
- Kolster, H., Janssens, T., Orban, G.A., Vanduffel, W., 2014. The retinotopic organization of macaque occipitotemporal cortex anterior to V4 and caudovernal to the middle temporal (MT) cluster. *J. Neurosci.* 34, 10168–10191.
- Konkle, T., Oliva, A., 2012. A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74, 1114–1124.
- Kraemer, P.M., Görner, M., Ramezani, H., Dicke, P.W., Thier, P., 2020. Frontal, parietal and temporal brain areas are differentially activated when disambiguating potential objects of joint attention. *eNeuro.*
- Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G., Mishkin, M., 2013. The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cognit. Sci.* 17, 26–49.
- Kustov, A.A., Robinson, D.L., 1996. Shared neural control of attentional shifts and eye movements. *Nature* 384, 74–77.
- Lafer-Sousa, R., Conway, B.R., 2013. Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. *Nat. Neurosci.* 16, 1870–1878.
- Laube, I., Kamphuis, S., Dicke, P.W., Thier, P., 2011. Cortical processing of head- and eye-gaze cues guiding joint social attention. *Neuroimage* 54, 1643–1653.
- Lee, D.K., Itti, L., Koch, C., Braun, J., 1999. Attention activates winner-take-all competition among visual filters. *Nat. Neurosci.* 2, 375–381.
- Lennert, T., Martinez-Trujillo, J., 2011. Strength of response suppression to distracter stimuli determines attentional-filtering performance in primate prefrontal neurons. *Neuron* 70, 141–152.
- Leopold, D.A., Krauzlis, R.J., 2020. How the brain pays attention to others' attention. *Proc. Natl. Acad. Sci. U.S.A.* 117, 3901–3903.
- Levy, I., Hasson, U., Avidan, G., Hendler, T., Malach, R., 2001. Center-periphery organization of human object areas. *Nat. Neurosci.* 4, 533–539.
- Li, Z., 2002. A saliency map in primary visual cortex. *Trends Cognit. Sci.* 6, 9–16.
- Logothetis, N.K., Pauls, J., Poggio, T., 1995. Shape representation in the inferior temporal cortex of monkeys. *Curr. Biol.* 5, 552–563.
- Lowet, E., Gomes, B., Srinivasan, K., Zhou, H., Schafer, R.J., Desimone, R., 2018. Enhanced neural processing by covert attention only during microsaccades directed toward the attended stimulus. *Neuron* 99, 207–214.e3.
- Mackey, W.E., Winawer, J., Curtis, C.E., 2017. Visual field map clusters in human frontoparietal cortex. *Elife* 6, e22974.
- Marciniak, K., Atabaki, A., Dicke, P.W., Thier, P., 2014. Disparate substrates for head gaze following and face perception in the monkey superior temporal sulcus. *Elife* 3.
- Marquardt, K., Ramezani, H., Dicke, P.W., Thier, P., 2017. Following eye gaze activates a patch in the posterior temporal cortex that is not part of the human "face patch" system. *eNeuro* 4.
- Mars, R.B., Sallet, J., Neubert, F.-X., Rushworth, M.F.S., 2013. Connectivity profiles reveal the relationship between brain areas for social cognition in human and monkey temporoparietal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 110, 10806–10811.
- Martinez-Trujillo, J.C., Treue, S., 2004. Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr. Biol.* 14, 744–751.
- Maunsell, J.H.R., Treue, S., 2006. Feature-based attention in visual cortex. *Trends Neurosci.* 29, 317–322.
- McPeck, R.M., Keller, E.L., 2002. Saccade target selection in the superior colliculus during a visual search task. *J. Neurophysiol.* 88, 2019–2034.
- Miller, E.K., Buschman, T.J., 2013. Cortical circuits for the control of attention. *Curr. Opin. Neurobiol.* 23, 216–222.
- Mishkin, M., Ungerleider, L.G., 1982. Contribution of striate inputs to the visuospatial functions of parieto-occipital cortex in monkeys. *Behav. Brain Res.* 6, 57–77.
- Mishkin, M., Ungerleider, L.G., Macko, K.A., 1983. Object vision and spatial vision: two cortical pathways. *Trends Neurosci.* 6, 414–417.
- Moore, T., Armstrong, K.M., 2003. Selective gating of visual signals by microstimulation of frontal cortex. *Nature* 421, 370–373.
- Moore, T., Fallah, M., 2001. Control of eye movements and spatial attention. *Proc. Natl. Acad. Sci. U.S.A.* 98, 1273–1276.
- Moore, T., Armstrong, K.M., Fallah, M., 2003. Visuomotor origins of covert spatial attention. *Neuron* 40, 671–683.
- Moran, J., Desimone, R., 1985. Selective attention gates visual processing in the extrastriate cortex. *Science* 229, 782–784.
- Motter, B.C., 1993. Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J. Neurophysiol.* 70, 909–919.
- Narumoto, J., Okada, T., Sadato, N., Fukui, K., Yonekura, Y., 2001. Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Brain Res. Cognit. Brain Res.* 12, 225–231.
- Nelissen, K., Vanduffel, W., Orban, G.A., 2006. Charting the lower superior temporal region, a new motion-sensitive region in monkey superior temporal sulcus. *J. Neurosci.* 26, 5929–5947.
- Nelissen, K., Borra, E., Gerbella, M., Rozzi, S., Luppino, G., Vanduffel, W., Rizzolatti, G., Orban, G.A., 2011. Action observation circuits in the macaque monkey cortex. *J. Neurosci.* 31, 3743–3756.
- Nobre, A.C., Gitelman, D.R., Dias, E.C., Mesulam, M.M., 2000. Covert visual spatial orienting and saccades: overlapping neural systems. *Neuroimage* 11, 210–216.
- Noudoost, B., Clark, K.L., Moore, T., 2014. A distinct contribution of the frontal eye field to the visual representation of saccadic targets. *J. Neurosci.* 34, 3687–3698.
- O'Connor, D.H., Fukui, M.M., Pinsk, M.A., Kastner, S., 2002. Attention modulates responses in the human lateral geniculate nucleus. *Nat. Neurosci.* 5, 1203–1209.
- Olshausen, B.A., Anderson, C.H., Essen, D.V., 1993. A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *J. Neurosci.* 13, 4700–4719.
- Ong, W.S., Madlon-Kay, S., Platt, M.L., 2020. Neuronal correlates of strategic cooperation in monkeys. *Nat. Neurosci.*
- Oristaglio, J., Schneider, D.M., Balan, P.F., Gottlieb, J., 2006. Integration of visuospatial and effector information during symbolically cued limb movements in monkey lateral intraparietal area. *J. Neurosci.* 26, 8310–8319.
- Patel, G.H., Yang, D., Jamerson, E.C., Snyder, L.H., Corbetta, M., Ferrera, V.P., 2015. Functional evolution of new and expanded attentional networks in humans. *Proc. Natl. Acad. Sci. U.S.A.* 112, 9454–9459.
- Pelphrey, K.A., Mitchell, T.V., McKeown, M.J., Goldstein, J., Allison, T., McCarthy, G., 2003. Brain activity evoked by the perception of human walking: controlling for meaningful coherent motion. *J. Neurosci.* 23, 6819–6825.
- Perani, D., Fazio, F., Borghese, N.A., Tettamanti, M., Ferrari, S., Decety, J., Gilardi, M.C., 2001. Different brain correlates for watching real and virtual hand actions. *Neuroimage* 14, 749–758.
- Perry, C.J., Fallah, M., 2014. Feature integration and object representations along the dorsal stream visual hierarchy. *Front. Comput. Neurosci.* 8.
- Petkov, C.I., Kayser, C., Augath, M., Logothetis, N.K., 2006. Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biol.* 4, e215.
- Pierno, A.C., Becchio, C., Wall, M.B., Smith, A.T., Castiello, U., 2006. Transfer of interfered motor patterns to self from others. *Eur. J. Neurosci.* 23, 1949–1955.
- Pitcher, D., Ungerleider, L.G., 2021. Evidence for a third visual pathway specialized for social perception. *Trends Cognit. Sci.* 25, 100–110.
- Pouget, A., Sejnowski, T.J., 1997. Spatial transformations in the parietal cortex using basis functions. *J. Cognit. Neurosci.* 9, 222–237.
- Ramezani, H., Thier, P., 2020. Decoding of the other's focus of attention by a temporal cortex module. *Proc. Natl. Acad. Sci. Unit. States Am.*
- Ramezani, H., Görner, M., Thier, P., 2021. Variability of neuronal responses in the posterior superior temporal sulcus predicts choice behavior during social interactions. *J. Neurophysiol.*
- Rees, G., Frackowiak, R., Frith, C., 1997. Two modulatory effects of attention that mediate object categorization in human cortex. *Science* 275, 835–838.
- Reynolds, J.H., Heeger, D.J., 2009. The normalization model of attention. *Neuron* 61, 168–185.
- Reynolds, J.H., Chelazzi, L., Desimone, R., 1999. Competitive mechanisms subserve attention in macaque areas V2 and V4. *J. Neurosci.* 19, 1736–1753.
- Richmond, B.J., Sato, T., 1987. Enhancement of inferior temporal neurons during visual discrimination. *J. Neurophysiol.* 58, 1292–1306.
- Rizzolatti, G., Craighero, L., 2004. The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.
- Rizzolatti, G., Sinigaglia, C., 2010. The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat. Rev. Neurosci.* 11, 264–274.
- Rizzolatti, G., Riggio, L., Dascola, I., Umiltà, C., 1987. Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia* 25, 31–40.
- Robinson, D.L., Petersen, S.E., 1992. The pulvinar and visual salience. *Trends Neurosci.* 15, 127–132.
- Rothblat, L., Pribram, K.H., 1972. Selective attention: input filter or response selection? An electrophysiological analysis. *Brain Res.* 39, 427–436.
- Roumazeilles, L., Schurz, M., Lojkiewicz, M., Verhagen, L., Schüfflgen, U., Marche, K., Mahmoodi, A., Emberton, A., Simpson, K., Joly, O., et al., 2021. Social prediction modulates activity of macaque superior temporal cortex. *Sci. Adv.* 7, eabh2392.
- Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y., Iwai, E., 1986. Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *J. Neurosci.* 6, 145–157.
- Sajad, A., Sadeh, M., Crawford, J.D., 2020. Spatiotemporal transformations for gaze control. *Physiol. Rep.* 8, e14533.
- Salinas, E., Abbott, L.F., 1997. Invariant visual responses from attentional gain fields. *J. Neurophysiol.* 77, 3267–3272.
- Salinas, E., Thier, P., 2000. Gain modulation: a major computational principle of the central nervous system. *Neuron* 27, 15–21.
- Sani, I., McPherson, B.C., Stemmann, H., Pestilli, F., Freiwald, W.A., 2019. Functionally defined white matter of the macaque monkey brain reveals a dorso-ventral attention network. *Elife* 8.
- Sani, I., Stemmann, H., Caron, B., Bullock, D., Stemmler, T., Fahle, M., Pestilli, F., Freiwald, W.A., 2021. The human endogenous attentional control network includes a ventro-temporal cortical node. *Nat. Commun.* 12, 360.
- Saxe, R., Kanwisher, N., 2003. People thinking about thinking people The role of the temporo-parietal junction in "theory of mind. *Neuroimage* 19, 1835–1842.

- Saygin, A.P., Sereno, M.I., 2008. Retinotopy and attention in human occipital, temporal, parietal, and frontal cortex. *Cerebr. Cortex* 18, 2158–2168.
- Schall, J.D., Hanes, D.P., 1993. Neural basis of saccade target selection in frontal eye field during visual search. *Nature* 366, 467–469.
- Schneider, W.X., 2013. Selective visual processing across competition episodes: a theory of task-driven visual attention and working memory. *Phil. Trans. Biol. Sci.* 368, 20130060.
- Seltzer, B., Pandya, D.N., 1989. Intrinsic connections and architectonics of the superior temporal sulcus in the rhesus monkey. *J. Comp. Neurol.* 290, 451–471.
- Seltzer, B., Pandya, D.N., 1994. Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. *J. Comp. Neurol.* 343, 445–463.
- Serences, J.T., Yantis, S., 2006. Selective visual attention and perceptual coherence. *Trends Cognit. Sci.* 10, 38–45.
- Sereno, A.B., Lehky, S.R., 2018. Attention effects on neural population representations for shape and location are stronger in the ventral than dorsal stream. *eNeuro* 5.
- Sereno, A.B., Lehky, S.R., Sereno, M.E., 2020. Representation of shape, space, and attention in monkey cortex. *Cortex* 122, 40–60.
- Smith, D.T., Schenk, T., 2012. The Premotor theory of attention: time to move on? *Neuropsychologia* 50, 1104–1114.
- Spitzer, H., Desimone, R., Moran, J., 1988. Increased attention enhances both behavioral and neuronal performance. *Science* 240, 338–340.
- Stemmann, H., Freiwald, W.A., 2016. Attentive motion discrimination recruits an area in inferotemporal cortex. *J. Neurosci. : Off. J. Soci. Neurosci.* 36, 11918–11928.
- Stemmann, H., Freiwald, W.A., 2019. Evidence for an attentional priority map in inferotemporal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 116, 23797–23805.
- Tanaka, K., 1996. Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* 19, 109–139.
- Thompson, K.G., Bichot, N.P., 2005. A visual salience map in the primate frontal eye field. *Prog. Brain Res.* 147, 251–262.
- Thompson, K.G., Bichot, N.P., Schall, J.D., 1997. Dissociation of visual discrimination from saccade programming in macaque frontal eye field. *J. Neurophysiol.* 77, 1046–1050.
- Thompson, J.C., Clarke, M., Stewart, T., Puce, A., 2005. Configural processing of biological motion in human superior temporal sulcus. *J. Neurosci.* 25, 9059–9066.
- Tolias, A.S., Sultan, F., Augath, M., Oeltermann, A., Tehovnik, E.J., Schiller, P.H., Logothetis, N.K., 2005. Mapping cortical activity elicited with electrical microstimulation using fMRI in the macaque. *Neuron* 48, 901–911.
- Treisman, A.M., Gelade, G., 1980. A feature-integration theory of attention. *Cognit. Psychol.* 12, 97–136.
- Treue, S., Maunsell, J.H.R., 1996. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* 382, 539–541.
- Tsao, D.Y., Freiwald, W.A., Tootell, R.B., Livingstone, M.S., 2006. A cortical region consisting entirely of face-selective cells. *Science (New York, NY)* 311, 670–674.
- Tsao, D.Y., Schweers, N., Moeller, S., Freiwald, W.A., 2008. Patches of face-selective cortex in the macaque frontal lobe. *Nat. Neurosci.* 11, 877–879.
- Tsotsos, J.K., 2011. *A Computational Perspective on Visual Attention*. MIT Press.
- Tsotsos, J.K., Kruijne, W., 2014. *Cognitive programs: software for attention's executive*. *Front. Psychol.* 5.
- Tsotsos, J.K., Culhane, S.M., Kei Wai, W.Y., Lai, Y., Davis, N., Nuflo, F., 1995. Modeling visual attention via selective tuning. *Artif. Intell.* 78, 507–545.
- Tsotsos, J.K., Abid, O., Kotseruba, I., Solbach, M.D., 2021. On the control of attentional processes in vision. *Cortex* 137, 305–329.
- Ullman, S., 1984. Visual routines. *Cognition* 18, 97–159.
- Vossel, S., Geng, J.J., Fink, G.R., 2014. Dorsal and ventral attention systems. *Neuroscientist* 20, 150–159.
- Wang, F., Chen, M., Yan, Y., Zhaoping, L., Li, W., 2015a. Modulation of neuronal responses by exogenous attention in macaque primary visual cortex. *J. Neurosci.* 35, 13419–13429.
- Wang, L., Mruczek, R.E.B., Arcaro, M.J., Kastner, S., 2015b. Probabilistic maps of visual topography in human cortex. *Cerebr. Cortex* 25, 3911–3931.
- Watanabe, J., Iwai, E., 1991. Neuronal activity in visual, auditory and polysensory areas in the monkey temporal cortex during visual fixation task. *Brain Res. Bull.* 26, 583–592.
- De Weerd, P., Desimone, R., Ungerleider, L.G., 2003a. Generalized deficits in visual selective attention after V4 and TEO lesions in macaques. *Eur. J. Neurosci.* 18, 1671–1691.
- De Weerd, P., Desimone, R., Ungerleider, L.G., 2003b. Impairments in spatial generalization of visual skills after V4 and TEO lesions in macaques (*Macaca mulatta*). *Behav. Neurosci.* 117, 1441–1447.
- Wilson, F.A., Scalaidhe, S.P., Goldman-Rakic, P.S., 1993. Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science* 260, 1955–1958.
- Wyk, B.C., Hudac, C.M., Carter, E.J., Sobel, D.M., Pelphrey, K.A., 2009. Action understanding in the superior temporal sulcus region. *Psychol. Sci.* 20, 771–777.
- Xu, R., Bichot, N.P., Takahashi, A., Desimone, R., 2021. The cortical connectome of primate lateral prefrontal cortex. *Neuron*.
- Yoo, S.-A., Tsotsos, J.K., Fallah, M., 2018. The attentional suppressive surround: eccentricity, location-based and feature-based effects and interactions. *Front. Neurosci.* 12.
- Yoshida, K., Saito, N., Iriki, A., Isoda, M., 2012. Social error monitoring in macaque frontal cortex. *Nat. Neurosci.* 15, 1307–1312.
- Zenon, A., Filali, N., Duhamel, J.-R., Olivier, E., 2010. Saliency representation in the parietal and frontal cortex. *J. Cognit. Neurosci.* 22, 918–930.